K-Winners-Take-All Computation with Neural Oscillators

Wei Wang and Jean-Jacques E. Slotine[1]

Nonlinear Systems Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts, 02139, USA
wangwei@mit.edu, jjs@mit.edu

Abstract

Artificial spike-based computation, inspired by models of computation in the central nervous system, may present significant performance advantages over traditional methods for specific types of large scale problems. This paper describes very simple network architectures for k-winners-take-all and soft-winner-take-all computation using neural oscillators. Fast convergence is achieved from arbitrary initial conditions, which makes the networks particularly suitable to track time-varying inputs.

1 Introduction

The discovery of synchronized oscillations in the visual cortex and other brain regions has triggered significant research in artificial spike-based computation [4, 9, 10, 14, 15, 17, 18, 21, 27, 29, 32, 33]. While neurons in the central nervous system are about six orders of magnitude "slower" than silicon-based elements, in both elementary computation time and signal transmission speed, their performance in networks often compares very favorably with their artificial counterparts even when reaction speed is concerned. In a sense, evolution may have been forced to develop extremely efficient computational schemes given available hardware limitations.

In a recent paper [34], we proposed new models for two common instances of such neural computation, winner-take-all and coincidence detection, featuring fast convergence and $O(n)$ network complexity. We saw that both computations could be achieved using a similar architecture, using global feedback inhibition in the first case, and global excitation in the second. In this paper, we further extend this computational architecture to k-winners-take-all and soft-winner-take-all.

Fast Winner-take-all (WTA) computation in [34] is based on the FitzHugh-Nagumo model, a well-known simplified version of the classical Hodgkin-Huxley model. Compared to previous

[1]To whom correspondence should be addressed.

WTA networks [1, 5, 11, 14, 16, 38], it has significant computational advantages. The network's initial states can be set arbitrarily, and convergence is guaranteed in at most two spiking periods, with a high computation resolution. The network's complexity is linear in the number of inputs and its size can be adjusted at any time during the computation. As this paper shows, by modifying the starting point of the global inhibitory neuron's charging mode, k-Winners-Take-All (k-WTA) can be computed instead. By running the charging mode independently, soft-Winner-Take-All (soft-WTA) can be computed. Both extensions inherit the advantages of the original WTA network.

After a brief review of the basic WTA network in section 2, k-WTA computation and soft-WTA computation are studied in Sections 3 and 4. Brief concluding remarks are offered in Section 5.

2 Winner-Take-All Network

The WTA network in [34] is based on the FitzHugh-Nagumo (FN) model [7, 24, 23]:

$$\begin{cases} \dot{v} = v(\alpha - v)(v - 1) - w + I \\ \dot{w} = \beta v - \gamma w \end{cases}$$

For appropriate parameter choices, there exists a unique equilibrium point for any given value of I, which is stable except for a finite range $I_l \leq I \leq I_h$ where the system tends to a limit cycle. The steady-state value of v at the stable equilibrium point increases with I.

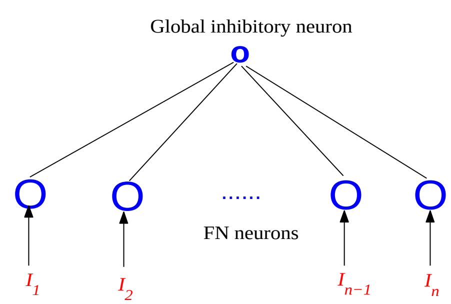

Figure 1: Diagram of the WTA network. There are n FN neurons receiving external inputs and a global inhibitory neuron monitoring the whole network.

The network structure is illustrated in Figure 1 where n FN neurons receive external stimulating inputs I_i and a global inhibition z. The dynamics of the FN neurons ($i = 1, \ldots, n$) are

$$\begin{cases} \dot{v}_i = v_i(\alpha - v_i)(v_i - 1) - w_i + I_i - z \\ \dot{w}_i = \beta v_i - \gamma w_i \end{cases}$$

The dynamics of the global inhibitory neuron is

$$\dot{z} = \begin{cases} -k_c\ (z - z_0) & \text{charging mode} \\ -k_d\ z & \text{discharging mode} \end{cases} \tag{1}$$

which starts charging if there is any FN neuron spiking in the network (i.e., if $\exists i,\ v_i \geq v_0$ for some given threshold v_0) and switches to discharging if the state z is saturated. With a fast charging rate k_c and a slow discharging rate k_d, the network computes the largest input (corresponding to the only spiking FN neuron) in at most two periods. Initial conditions can be set arbitrarily and the computation resolution is very high. Detailed analysis and discussions can be found in [34].

3 K-Winners-Take-All Network

K-WTA, a common variation of WTA computation where the output indicates for each neuron whether its input is among the k largest, has been studied in such fields as competitive learning, pattern recognition and pattern classification [3, 8, 31, 36, 37]. As Maass argued in [19], in principle a k-WTA network can replace a two-layer threshold circuit to perform most standard nonlinear computational operations.

Most k-WTA studies are based on steady-state stability analysis. Many models define the winners as the neurons with the largest initial states [22, 36] or require initial conditions to be set precisely [37], making the networks not well suited to time-varying inputs. Others adopt particular design methodologies [25, 26] but the network size or the number of winners is limited. K-WTA is also implemented in analog VLSI circuits [31], which extend the elegant WTA model in [16] but inherit its low resolution limit as well.

The neural network described in Section 2 can be easily extended to k-WTA computation, where an FN neuron spikes if and only if its input is among the k largest. Indeed, as the global inhibition force decreases, the FN neurons enter the oscillation region *rank-ordered* by their inputs. Thus, while for WTA computation, the global inhibition neuron is charged after the first arrival, for k-WTA computation the charging moment is simply modified to capture the k^{th} arrival instead.

To this effect, we augment the dynamics of the FN neuron with an additional state variable u_i (for simplicity, we shall still call FN neuron such a generalized element)

$$\begin{cases} \dot{v}_i = v_i(\alpha - v_i)(v_i - 1) - w_i + I_i - u_i - z \\ \dot{w}_i = \beta v_i - \gamma w_i \\ \dot{u}_i = k_u\ (\zeta_i u_0 - u_i) \end{cases}$$

where u_0 is a constant saturation value and k_u the charging/discharging rate. The variable ζ_i takes two values, namely it switches to 0 whenever z approaches a saturation value z_0,

else it switches to 1 if v_i exceeds a given threshold v_0. This make the dynamics of u_i a local self-inhibition, which starts charging if the basic FN neuron spikes and discharges whenever the global inhibitory neuron spikes. Note that the specific form of the dynamics of u_i can be more general, as long as the value of u_i varies between 0 and u_0, and the transition periods are very fast (which is satisfied here by choosing a large k_u).

The dynamics of the global inhibitory neuron is the same as (1), except that we start its charging mode if any k FN neurons in the network spike. Such a moment can be captured by determining that $\sum_{i=1}^{n} u_i$ approaches ku_0. Thus, if any FN neuron spikes, it excites only the corresponding local inhibitory portion but has no effect on the rest of the network. If there are k local inhibitions turned on, the global inhibitory neuron is charged, which then releases all the local inhibitions and starts a new period.

Compared to the WTA network in Section 2, the basic principle underlying the k-WTA network described above is the same, exploiting the simple properties of the FN model. Thus, most of the computational advantages of the WTA network [34] are inherited by the k-WTA extension. In particular

- The initial conditions of the network can be set arbitrarily.
- With appropriate parameters, the computation can be completed at most in two periods, where the first period is affected by the initial conditions but the k spiking neurons during the following periods are guaranteed to be those with the largest inputs. If the initial inhibitions are strong, the computation is completed in one period.
- Since initial conditions are immaterial and the computation speed is very fast, the k-WTA network is able to track time-varying inputs. Moreover, since the network complexity is $O(n)$, individual FN neurons can be added or removed at any time during the computation.
- The inputs I_i should be lower-bounded by I_l, the lower threshold of the FN oscillation region. It should also be upper-bounded to set inhibition saturations, although the upper bound value is not restricted.
- The computation resolution also follows that of the WTA network. It can be improved by decreasing the discharging rate k_d, as well as the relaxation time of the FN neurons.
- FN neurons receiving equal inputs behave identically, which means that the k-WTA computation may generate more than k winners in this particular case.

Example 3.1: The result is illustrated in simulation in Figure 2, with $n = 10$ and $k = 3$. The parameters of the FN neurons are set as $\alpha = 5.32, \beta = 3, \gamma = 0.1$, with spiking threshold $v_0 = 5$. The parameters of the local inhibition are $u_0 = 160, k_u = 100$. The inputs I_i are chosen randomly from 20 to 125. The parameters of the global neuron are $z_0 = 240, k_c = 100, k_d = 1/40$. All initial conditions are chosen arbitrarily. The three spiking neurons after the first charging of the global neuron are those with the three largest inputs.

Note that the output frequency is determined mainly by the global neuron's dynamics and the value of the k^{th} largest input. It can be increased by increasing the global neuron's discharging rate after the first winner spikes so as to facilitate the other winners' spiking. $\square$

Example 3.2: Figure 3 illustrates a simulation result with $n = 3$ and $k = 2$. The parameters are the same as those in Example 3.1. The inputs keep varying and switch winning positions several times. The spiking neurons always track the two largest inputs. $\square$

4 Soft-Winner-Take-All

Soft-WTA [19] (or softmax) is another variation of WTA computation, where the outputs reflect the rank of all inputs according to their size. Although soft-WTA is a very powerful primitive [19, 20] in that it can be used to compute any continuous function, its "neural" implementation is complex. Recently, [39] studied soft-WTA as an optimization problem not based on a biologically plausible mechanism; [13] presented a hardware model of selective visual attention which lets the attention switch between the selected inputs, but whose switching order does not completely reflect the input ranks. In this section, we develop a simple neural network which computes soft-WTA very fast and generates spiking outputs rank-ordered by their inputs.

Let $k = n$ in the k-WTA network described in Section 3. Then we get a pre-ordered spiking sequence in each stable period, since all the FN neurons enter the oscillation region and then spike rank-ordered by their inputs. However, such an n^{th} arrival moment may not be measurable if the number of inputs n is unknown or time-varying. To avoid this problem and make the solution more general, we let the charging mode of the global inhibitory neuron start only if the inhibition z is lower than a given bound z_{low}. The spikings of all the FN neurons in the network are guaranteed by the condition that

$$z_{low} \ < \ I_l + I_{min}$$

where I_l is the lower bound of the oscillation region of the FN model and I_{min} is the minimum input value.

Example 4.1: Figure 4 illustrates the result in simulation with $n = 10$. The parameters are the same as those in Example 3.1. The inputs I_i are distributed uniformly between 80 and 120. The inhibition lower bound is $z_{low} = 60$. Initial conditions are chosen arbitrarily. The computation is completed in the second period, during and after which the spiking times of the FN neurons are ranked by their inputs. $\square$

The simple soft-WTA network presented above inherits the main computational advantages of our WTA and k-WTA networks. Initial conditions can be arbitrary, the computation

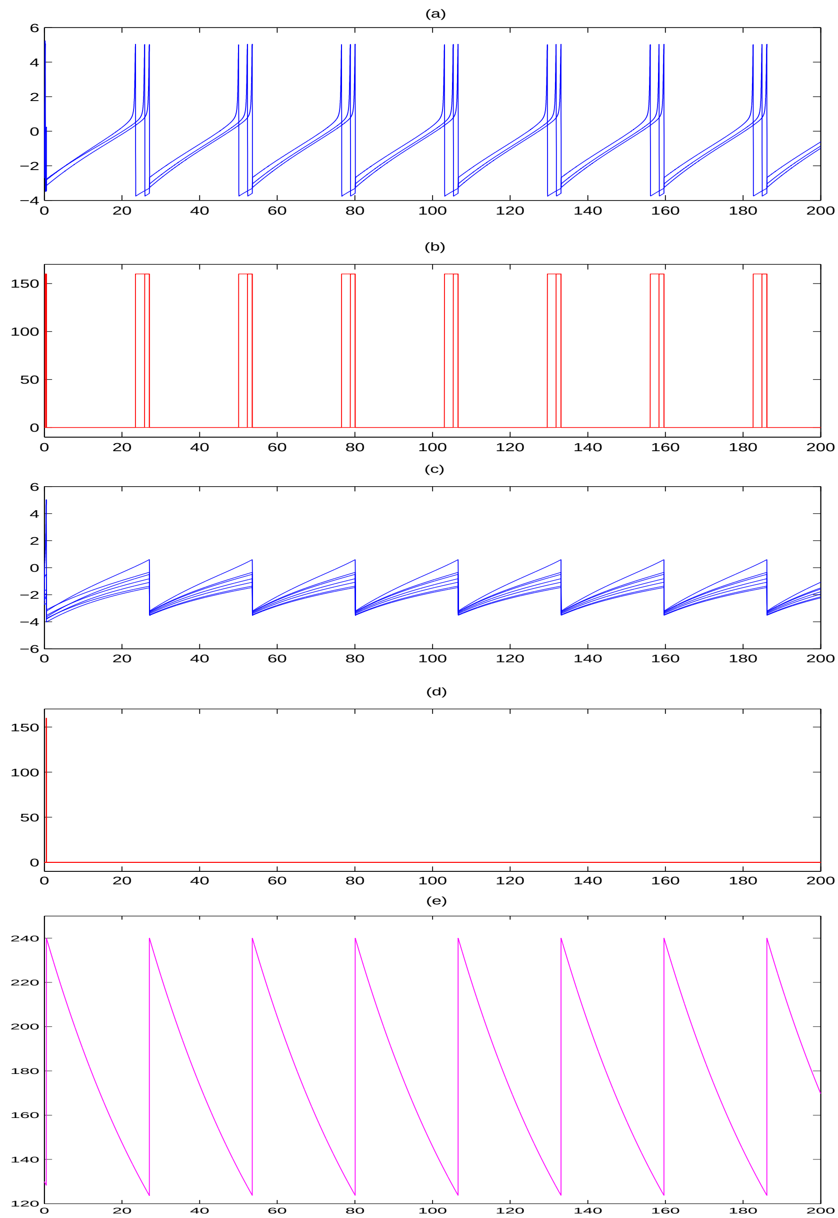

Figure 2: k-WTA computation result of Example 3.1 with $n = 10$ and $k = 3$. The plots are the time developments of (a) v_i of the neurons with the three largest inputs; (b) u_i of the neurons with the three largest inputs; (c) v_i of the other seven neurons; (d) u_i of the other seven neurons; (e) global inhibition z. The computation is completed in less than two periods.

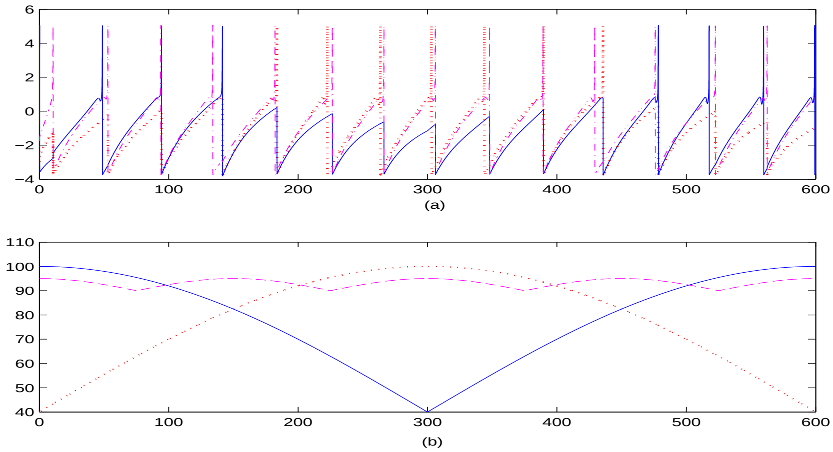

Figure 3: k-WTA computation result of Example 3.2 with $n = 3$ and $k = 2$. The inputs are not constant. The spiking neurons always track the two largest inputs. The plots are (a).states v_i versus time; (b).inputs I_i versus time. Note that each v_i in plot (a) corresponds to the I_i in plot (b) with the same line type (solid, dashed or dotted).

is completed in at most two periods, network complexity is linear and neurons can be added or removed at any time. We expect it to be effective in many applications such as selective attention, associative memory and competitive learning, and also to provide an efficient desynchronization mechanism for perceptional binding [10, 27, 32, 33].

5 Concluding Remarks

Basic neural computations such as winner-take-all, k-winners-take-all, soft-winner-take-all, and coincidence detection can all be implemented using a common architecture and biological plausible neuron models. Fast and robust convergence is guaranteed, and time-varying inputs can be tracked. Further research will study models of higher level brain functions, such as perception, based on the WTA networks and on general nonlinear synchronization mechanisms derived in [28, 35].

Acknowledgments: This work was supported in part by a grant from the National Institutes of Health. The authors benefited from stimulating discussions with Matthew Tresch.

References

[1] Arbib, M.A. (1995) The Handbook of Brain Theory and Neural Networks, *MIT Press*(Cambridge, MA, USA)

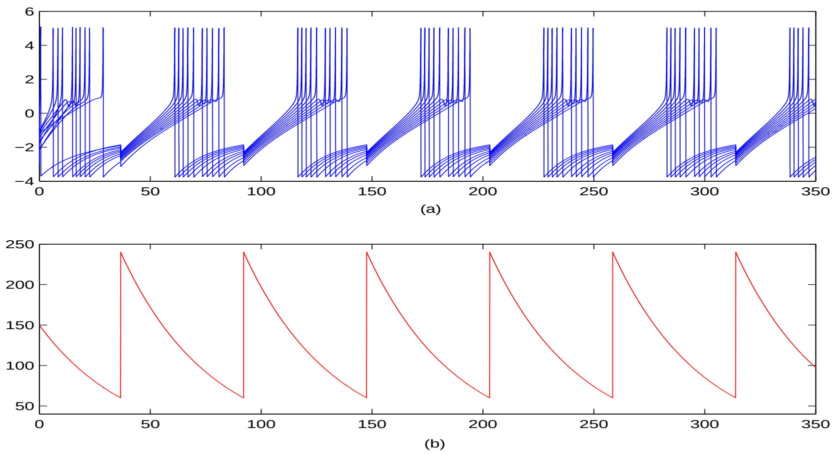

Figure 4: Soft-WTA computation result of Example 4.1 with $n = 10$. The plots are (a).states v_i versus time; (b).global inhibition z versus time. The initial conditions are chosen arbitrarily and the computation is completed in the second period.

[2] Amari, S., and Arbib, M. (1977) Competition and Cooperation in Neural Nets, in *Systems Neuroscience*, J.Metzler, Ed., Academic Press (San Diego) 119-165

[3] Badel, S., Schmid, A., and Leblebici Y. (2003) A VLSI Hamming Artificial Neural Network with k-Winner-Take-All and k-Loser-Take-All Capability, *International Joint Conference on Neural Networks*, Portland, OR, 977-982

[4] Brody, C.D., and Hopfield, J.J. (2003) Simple Networks for Spike-Timing-Based Computation, with Application to Olfactory Processing, *Neuron*, 37:843-852

[5] Fang, Y., Cohen, M., and Kincaid, M. (1996) Dynamics of a Winner-Take-All Neural Network, *Neural Networks*, 9:7, 1141-1154

[6] Feldman, J.A., and Ballard, D.H. (1982) Connectionist Models and Their Properties, *Cognit. Sci.*, vol.6:205-254

[7] FitzHugh, R.A. (1961) Impulses and Physiological States in Theoretical Models of Nerve Membrane, *Biophys. J.*, 1:445-466

[8] Fukai, T., and Tanaka, S. (1997) A Simple Neural Network Exhibiting Selective Activation of Neuronal Ensembles: From Winner-Take-All to Winners-Share-All, *Neural Computation*, 9:77-97

[9] Gerstner, W. (2001) A Framework for Spiking Neuron Models: The Spike Response Model, In *The Handbook of Biological Physics*, vol.4:469-516

[10] Gray, C.M. (1999) The Temporal Correlation Hypothesis of Visual Feature Integration: Still Alive and Well, *Neuron*, 24:31-47

[11] Grossberg, S. (1973) Contour Enhancement, Short-Term Memory, and Constancies in Reverberating Neural Networks, *Studies in Applied Mathematics*, 52:217-257

[12] Grossberg, S. (1978) Competition, Decision, and Consensus, *Journal of Mathematical Analysis and Applications*, 66:470-493

[13] Indiveri, G. (2000) A 2D Neuromorphic VLSI Architecture for Modeling Selective Attention, *IEEE-INNS-ENNS International Joint Conference on Neural Networks*, vol.4, July 24-27, Como, Italy

[14] Jin, D.Z., and Seung, H.S. (2002) Fast Computation With Spikes in a Recurrent Neural Network, *Physical Review E*, 65:051922

[15] Kazantsev, V.B., Nekorkin, V.I., Makarenko, V.I., and Llinás, R. (2003) Olivo-Cerebellar Cluster-Based Universal Control System, *PNAS*, vol.100 no.22:13064-13068

[16] Lazzaro, J., Ryckebusch, S., Mahowald, M.A., and Mead, C. (1988) Winner-Take-All Networks of $O(n)$ Complexity, in *Advances in Neural Information Processing Systems 1.* San Mateo, CA:Morgan Kaufmann Publishers, 703-711

[17] Llinás, R., Ribary, U., Contreras, D., and Pedroarena, C. (1998) The Neuronal Basis for Consciousness, *Phil. Trans. R. Soc. Lond.* B 353:1841-1849

[18] Llinás, R.R., Leznik, E., and Urbano, F.J. (2002) Temporal Binding via Cortical Coincidence Detection of Specific and Nonspecific Thalamocortical Inputs: a Voltage-Dependent Dye-Imaging Study in Mouse Brain Slices, *PNAS*, vol.99 no.1:449-454

[19] Maass, W. (2000) On the Computational Power of Winner-Take-All, *Neural Comput.*, 12(11):2519-35

[20] Maass, W. (2000) Neural Computation with Winner-Take-All as The Only Nonlinear Operation, in *Advances in Information Processing Systems*, vol.12:293-299, MIT Press

[21] Maass, W., Natschläger, T., and Markram, H. (2003) On the Computational Power of Circuits of Spiking Neurons, submitted

[22] Majani, E., Erlanson, R., and Abu-Mostafa, Y. (1989) On the k-Winners-Take-All Network, in *Advances in Neural Information Processing Systems I*, D.S. Touretzky, Ed. 634-642

[23] Murray, J.D. (1993) Mathematical Biology, *Berlin;New York: Springer-Verlag*

[24] Nagumo, J., Arimoto, S., and Yoshizawa, S. (1962) An Active Pulse Transmission Line Simulating Nerve Axon, *Proc. Inst. Radio Engineers*, 50:2061-2070

[25] Perfetti, R. (1995) On the Robust Design of k-Winners-Take-All Networks, *IEEE Transactions on Circuits and Systems-II*, 42(1):55-58

[26] Seiler, G., and Nossek, J.A. (1993) Winner-Take-All Cellular Neural Networks, *IEEE Trans. Circuits Syst.-II*, 40(3):184-190

[27] Singer, W., and Gray, C.M. (1995) Visual Feature Integration and The Temporal Correlation Hypothesis, *Annu. Rev. Neurosci.*, 18:555-586

[28] Slotine, J.J.E., and Wang, W. (2003) A Study of Synchronization and Group Cooperation Using Partial Contraction Theory, in *Cooperative Control*, Kumar V. Editor, Springer-Verlag

[29] Thorpe, S., Delorme, A., and Rullen, R.V. (2001) Spike-Based Strategies for Rapid Processing, *Neural Networks*, 14:715-725

[30] Tononi, G., Srinivasan, R., Russell, D.P., and Edelman G.M. (1998) Investigating Neural Correlates of Conscious Perception by Frequency-Tagged Neuromagnetic Responses, *PNAS* 95:3198-3203

[31] Urahama, K., and Nagao, T. (1995) K-Winner-Take-All Circuit with $O(N)$ Complexity, *IEEE Trans. on Neural Networks*, vol.6:776-778

[32] von der Malsburg, C. (1995) Binding in Models of Perception and Brain Function, *Current Opinion in Neurobiology*, 5:520-526

[33] Wang, D.L. (2002) The Time Dimension for Neural Computation, *Technical Report OSU-CISRC-09/02-TR21*, Department of Computer and Information Science, The Ohio State University

[34] Wang, W., and Slotine, J.J.E. (2003) Fast Computation with Neural Oscillators, submitted to *Physical Review Letters*, also available at *e-Print archive* http://www.arxiv.org/PS_cache/q-bio/pdf/0312/0312025.pdf

[35] Wang, W., and Slotine, J.J.E. (2003) On Partial Contraction Analysis for Coupled Nonlinear Oscillators, submitted to *Biological Cybernetics*

[36] Wolfe, W.J., *et al.* (1991) K-Winner Networks, *IEEE Transactions on Neural Networks*, 2(2):310-315

[37] Yen, J.C., Guo, J.I., and Chen, H.C. (1998) A New k-Winners-Take-All Neural Network and Its Array Architecture, *IEEE Transactions on Neural Networks*, 9(5):901-912

[38] Yuille, A.L., and Grzywacz, N.M. (1989) A Winner-Take-All Mechanism Based on Presynaptic Inhibition, *Neural Computation*, 1:334-347

[39] Yuille, A.L., and Geiger, D. (2002) Winner-Take-All Networks, in *The Handbook of Brain Theory and Neural Networks*, 2nd ed., Michael A. Arbib, Editor, The MIT Press, 1228-1231